

Application of Spectral Accuracy to Improve the Identification of Organic Compounds in Environmental Analysis

Emmanuel Eysseric,[†] Killian Barry,[†] Francis Beaudry,[‡] Magali Houde,[§] Christian Gagnon,[§] and Pedro A. Segura^{*,†,‡,§}

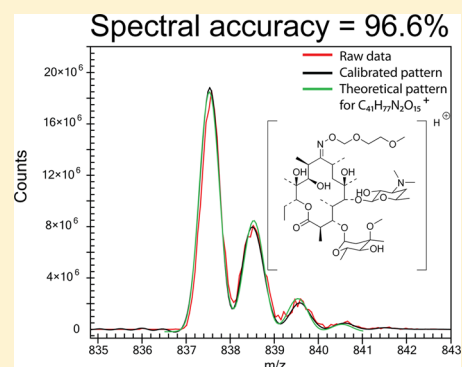
[†]Department of Chemistry, Université de Sherbrooke, Sherbrooke, Quebec J1K 2R1, Canada

[‡]Groupe de Recherche en Pharmacologie Animal du Québec (GREPAQ), Department of Veterinary Biomedicine, Université de Montréal, Saint-Hyacinthe, Quebec J2S 2M2, Canada

[§]Environment and Climate Change Canada, Montreal, Quebec H2Y 2E7, Canada

S Supporting Information

ABSTRACT: Correct identification of a chemical substance in environmental samples based only on accurate mass measurements can be difficult especially for molecules >300 Da. Here is presented the application of spectral accuracy, a tool for the comparison of isotope patterns toward molecular formula generation, as a complementary technique to assist in the identification process of organic micropollutants and their transformation products in surface water. A set of nine common contaminants (five antibiotics, an herbicide, a beta-blocker, an anti-depressant, and an antineoplastic) frequently found in surface water were spiked in methanol and surface water extracts at two different concentrations (80 and 300 $\mu\text{g L}^{-1}$). They were then injected into three different mass analyzers (triple quadrupole, quadrupole-time-of-flight, and quadrupole-orbitrap) to study the impact of matrix composition, analyte concentration, and mass resolution on the correct identification of molecular formulas using spectral accuracy. High spectral accuracy and ranking of the correct molecular formula were in many cases compound-specific due principally to conditions affecting signal intensity such as matrix effects and concentration. However, in general, results showed that higher concentrations and higher resolutions favored ranking the correct formula in the top 10. Using spectral accuracy and mass accuracy it was possible to reduce the number of possible molecular formulas for organic compounds of relative high molecular mass (e.g., between 400 and 900 Da) to less than 10 and in some cases, it was possible to unambiguously assign one specific molecular formula to an experimental isotopic pattern. This study confirmed that spectral accuracy can be used as a complementary diagnostic technique to improve confidence levels for the identification of organic contaminants under environmental conditions.



The identification of organic micropollutants, such as pesticides, pharmaceuticals, personal care products, plastic additives, and their metabolites, is a real challenge as a large number and variety of compounds are present in the environment.¹ One of the first steps to identify a compound using mass spectrometry (MS) is to determine the molecular formula from its mass spectrum. Recently, Schymanski et al.² proposed to improve the communication of identification confidence using MS based on a five-level approach. According to the authors, accurate mass represents the lowest confidence (level 5), followed by unequivocal molecular formula (level 4), tentative candidate (level 3, based for example on tandem mass spectrometry or other experimental data), probable structure (level 2, which could be reached using library spectrum match or other diagnostic evidence), and finally confirmed structure (level 1, which requires a reference standard). Determination of an unequivocal chemical formula (level 4) with mass accuracy <5 ppm for a unknown peak in the mass spectrum is challenging for compounds with molecular masses >300 Da³ and can often lead to incorrect conclusions. Using tandem mass spectrometry (MS/MS) databases, such as mzCloud or Mass

Bank, may help to reach confidence level 2 by searching library spectrum matches, however it is difficult to perform such confirmation for less known organic micropollutants or their transformation products that might be absent from those databases. For that reason, complementary techniques have been recently developed to obtain more information on the composition and structure of unknowns such as postcolumn hydrogen–deuterium exchange⁴ and comparison of data acquired in the positive and negative ionization mode.^{5,6} The former could be used as diagnostic evidence to obtain more information about the presence of specific functional groups (e.g., exchangeable hydrogens present in alcohol and amine groups) on a molecule, while the latter can be employed to confirm the presence of certain compounds. Dual ionization has been used for metabolomic profiling in the past to broaden the range of detection of MS methods. Additionally, for some environmental

Received: May 10, 2017

Accepted: August 2, 2017

Published: August 2, 2017

contaminants, such as pharmaceuticals, herbicides, and fluorinated compounds, electrospray in the negative mode is the preferred ionization polarity.⁷ Obtaining additional structural information from the mass spectrum for a precursor ion before carrying on MS/MS experiments is of interest since it could save time and resources. It could also help in the identification process of unknown organic micropollutants and their transformation products, often present in samples at low concentrations, since it is difficult or even impossible to obtain meaningful MS/MS spectra for peaks of low abundance.

Previous studies have explored different approaches to improve the identification of small organic compounds using MS. The “Seven Golden Rules” established by Kind and Fiehn⁸ provide a way to limit the sheer number of possible formulas and are now widely used in the identification process of compounds. In these rules, isotope pattern is the major component for the formula determination, along with other rules, such as hydrogen to carbon ratios and probable elements. Recent studies based exclusively on accurate mass to assess molecular formula are rare because MS/MS is normally used to provide additional structural information; however alternatives have been explored. Garcia-Reyes et al.⁹ previously developed a workflow for detecting and identifying pesticides and their degradation products using liquid chromatography-time-of-flight MS. The proposed method is efficient for compounds containing S, Cl, or Br, which are a common occurrence in pesticides, because of their very easily distinguishable isotopic patterns. However, if an unknown compound does not contain such elements, the number of possible molecular formulas for a given accurate mass within an acceptable tolerance window cannot be reduced significantly, which impairs the compound identification process. For example, 187 possible molecular formulas, within 5 mDa of mass error and having between 0 to 50 atoms of C, H, N, O, P, and F with up to 20 double bond equivalents, were found for an hypothetical ion of m/z 400.1234. Another approach proposed by Little et al.,^{10,11} used accurate mass to perform searches on databases such as the Chemical Abstracts Service or ChemSpider to identify unknown compounds. The authors applied orthogonal filters based on the number of literature references to prioritize the list of potential candidates. Though this approach can be highly successful to identify compounds that are commercially available, it can be less suitable for the identification of transformation products of environmental contaminants that may not be integrated in databases.

An interesting alternative technique to improve the confidence on the identification level of a given unknown is spectral accuracy for MS as introduced by Wang and Gu.¹² Spectral accuracy is a metric that describes the similarity between a calibrated experimental profile data of an ion, obtained through a mathematical transformation of its experimental profile data and the theoretical (calculated) isotopic pattern corresponding to a given molecular formula. Thus, high spectral accuracy (e.g., > 98%) indicates that the experimental isotopic pattern fits well to the isotopic pattern of a specific molecular formula within 2% spectral error¹² (Figure 1). The main advantage of spectral accuracy over mass accuracy is that in the latter error is measured at a single point, while in the former error is measured throughout the whole isotopic distribution. Therefore, spectral accuracy uses all the information embedded in the experimental mass spectrum to rank possible molecular formulas accordingly to their level of likeness to the theoretical mass spectrum. Spectral accuracy was successfully applied to

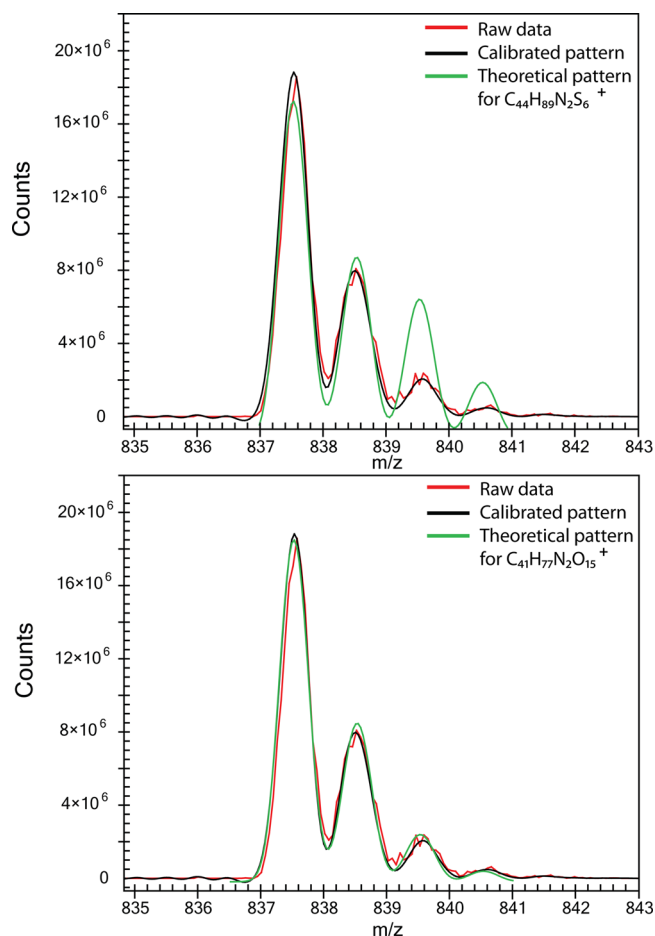


Figure 1. Determination of spectral accuracy in the MassWorks software using a known compound spiked in MeOH and analyzed by LC-ESI(+)-QqQMS. The raw data is first calibrated, that is, mathematically transformed, and then compared to the theoretical isotopic pattern of molecular formula candidates. The top figure shows that the isotopic pattern corresponding to an ion of formula $C_{44}H_{89}N_2S_6^+$ ($\Delta m = -3.4$ mDa), is not a good match for the calibrated isotopic pattern, therefore the spectral accuracy is 75.00% and was ranked 71st out of 71 possible formulas. The bottom figure shows a better match between the calibrated and the theoretical isotopic pattern of ion $C_{41}H_{77}N_2O_{15}^+$ ($\Delta m = -0.7$ mDa) and the spectral accuracy (96.63%), was ranked 6th out of 71 possible for molecular formulas within the software constraints (C_{1-66} , H_{0-109} , N_{0-20} , O_{0-25} , S_{0-13} , mass tolerance = 5 mDa, electron state: even, double bond equivalents range = -0.5 to 50). That ion corresponds to the protonated molecule of roxithromycin, the compound spiked in the sample.

the identification of high mass (639–1664 Da) natural products in a previous study using a linear ion trap-orbitrap mass spectrometer.¹³ On the basis of both mass and spectral accuracy, it was possible in some cases to eliminate >99% of formula candidates and the correct formula was usually ranked among the top five candidates. However, data were obtained using millimolar concentrations in pure solvents, experimental conditions that do not correspond to environmental analysis. Moreover, previous works on spectral accuracy showed that ion abundance was a major factor impacting quality of results.¹⁴

The main objective of the present work was therefore to determine if spectral accuracy could be used as a complementary technique for the identification of organic contaminants in environmental sample analysis. To answer this

question, a set of frequently detected organic contaminants were spiked in methanol and surface water extracts at two different concentrations (80 and 300 $\mu\text{g L}^{-1}$). Three different types of mass analyzers (i.e., triple quadrupole, quadrupole-time-of-flight, and quadrupole-orbital ion trap) were used to study the impacts of three important factors in environmental analysis (matrix composition, analyte concentration and mass resolution) on spectral accuracy and formula ranking of the selected organic micropollutants (Figure SI-1).

EXPERIMENTAL SECTION

Reagents and Standards. Water, acetonitrile (ACN), methanol (MeOH), and 0.1% formic acid (FA) in ACN were all LC-MS Optima grade and were obtained from Fisher Scientific (Canada). Analytical standards of atrazine (acronym, ATZ; purity, 98.1%), fluoxetine hydrochloride (FLX, 99.95%), josamycin (JOS, $\geq 98\%$), metoprolol tartrate (MTP, $\geq 98\%$), ofloxacin (OFL, 99.8%) roxithromycin (ROX, $\geq 90\%$), and sulfamethoxazole (SMX, 99.9%) were purchased from Sigma-Aldrich (Canada). Trimethoprim (TRI, $>98\%$) and methotrexate (MTX, $\geq 98\%$) were purchased from Santa Cruz Biotechnology (USA).

Collection and Preparation of Samples. Mixture solutions of the aforementioned compounds were prepared in MeOH each at high (300 $\mu\text{g L}^{-1}$) and low concentration (80 $\mu\text{g L}^{-1}$). River surface water samples (250 mL) were collected from the Magog River (Sherbrooke, Quebec) on September 13, 2016 in amber high density polyethylene bottles and conserved in an ice cooler until arrival to the laboratory, where they were immediately stored at $-20\text{ }^{\circ}\text{C}$. Upon extraction, samples were thawed at room temperature and buffered to pH 7 with phosphoric acid monobasic and phosphoric acid dibasic. Water samples were then extracted using a previously reported method.⁴ Briefly, samples were loaded in polymeric Strata-X solid-phase extraction cartridges (Phenomenex, USA) and then eluted with $2 \times 3\text{ mL}$ of an ACN-MeOH (1:1, v/v) solution. Eluates were evaporated under a nitrogen flow and reconstituted to 10 mL with MeOH spiked at the same concentrations as the previous solutions (80 and 300 $\mu\text{g L}^{-1}$). Therefore, considering a preconcentration factor of 25 (initial volume = 250 mL, final volume = 10 mL), the high and low concentrations used for each compound are equivalent to 3 and 12 $\mu\text{g L}^{-1}$ respectively, which are in the high range of environmental concentrations of many organic micropollutants detected in surface waters around the globe.^{15,16} For all experiments, samples were injected three times in order to evaluate reproducibility of the results.

Another set of analysis at lower concentrations (10, 20, 30, 40, and 50 $\mu\text{g L}^{-1}$) was added in order to assess the low concentration limits of spectral accuracy on the quadrupole-time-of-flight mass spectrometer and the quadrupole-orbital ion trap mass spectrometer. Considering the preconcentration factor, those concentrations were equivalent to 0.38, 0.75, 1.1, 1.5, and 1.9 $\mu\text{g L}^{-1}$. For these experiments, the same sample was injected two times.

Instruments and Methods. *Liquid Chromatography-Triple Quadrupole Mass Spectrometry (LC-QqQMS).* The liquid chromatography-triple quadrupole mass spectrometry instrument (LC-QqQMS) used for this work involved an Acquity Ultra Performance LC coupled to a Quattro Premier XE triple quadrupole mass spectrometer both manufactured by Waters (USA). The LC column was a Waters Acquity UPLC HSS T3 $2.1 \times 50\text{ mm}$, $1.8\text{ }\mu\text{m}$. The mobile phase was

composed of solvent A, water with 0.1% (v/v) formic acid, while solvent B was acetonitrile with 0.1% (v/v) formic acid. Chromatographic separation was obtained using the following elution gradient: at initial time, 5% B; at 5 min, 100% B; at 7 min, 100% B; at 7.01 min, 5% B; at 10 min, 5% B. Run time was 10 min. Mobile phase flow rate was $500\text{ }\mu\text{L min}^{-1}$ throughout the run. Injection volume was $10\text{ }\mu\text{L}$. The QqQMS system was first mass calibrated via MassLynx with a solution of sodium formate before the acquisition. Additional mass calibration was performed through MassWorks by infusing an external calibrant (sodium formate) before the acquisition or by LC-QqQMS analysis of internal calibrants (MTP, FLX, OFL and JOS) mixed with the samples. A detailed description of experiments carried out to evaluate the mass accuracy stability of the QqQMS system is presented in section SI-1 of the Supporting Information.

Liquid Chromatography-Quadrupole-Time-of-Flight Mass Spectrometry (LC-QqTOF-MS). The instrument used in these experiments involved a LC system manufactured by Shimadzu (Japan) and composed of a Nexera LC-30AD pump module, a SIL-30AC autosampler and a CTO-30A as column oven module. This LC system was coupled to a high-resolution mass spectrometer, the Maxis quadrupole-time-of-flight mass spectrometer (QqTOFMS) made by Bruker (USA). LC conditions were identical as those used in the LC-QqQMS setup. Injection volume was $1\text{ }\mu\text{L}$. The QqTOFMS was calibrated with a sodium formate solution in HPC mode after waiting 30 min for the system to stabilize. The mass drift was monitored and all analyses were done within 4 h of the calibration. No lock mass solution was used. In these conditions, full width at half-maximum mass resolution (R_{FWHM}) at m/z 455 was about 25 000.

Liquid Chromatography-Quadrupole-Orbital Ion Trap Mass Spectrometry (LC-QqOrbitrap MS). The instrument used in these experiments was composed of a Dionex Ultimate 3000 Rapid Separation System ultrahigh performance liquid chromatograph coupled to a Q-Exactive hybrid quadrupole-orbital ion trap mass spectrometer (QqOrbitrapMS) both manufactured by Thermo (USA). LC conditions were the same as for the LC-QqQMS setup. Injection volume was $2\text{ }\mu\text{L}$. As for the QqOrbitrapMS parameters, the ion source was ESI in positive mode, spray voltage was at 4.2 kV, capillary temperature was at $300\text{ }^{\circ}\text{C}$, sheath gas was 50 arbitrary units and auxiliary gas was 25 arbitrary units. R_{FWHM} was set to 70 000 or 140 000 at m/z 400. Mass range was m/z 200–1000. The instrument was calibrated using the calibrant solution recommended by the manufacturer, a solution containing *n*-butylamine, caffeine, the tripeptide MRFA and Ultramark 1621, a mixture of fluorinated phosphazines.

Software Parameters. Raw profile data of the mass spectrum of the test analytes spiked in MeOH and river water samples were extracted from acquisition files and processed in MassWorks (Cerno Bioscience, USA). This software uses two algorithms for the determination of spectral accuracy and formula ranking: calibrated line shape isotope profile search (CLIPS) and self-calibrated line shape isotope profile search (sCLIPS). The former is used for low resolution data and the latter for high-resolution data. Parameters for both algorithms are shown on Table 1. As the analyses were performed with electrospray ionization in the positive ionization mode, the observed charges were mainly +1 ($[M + H]^+$). Doubly charged species were only observed for ROX, albeit it was not the major ion. Mass tolerance was highly dependent on the type of calibration used for the CLIPS algorithm. For the internal calibration with CLIPS, it was

Table 1. Parameters for MassWorks Algorithms

algorithm	charge	mass tolerance (ppm)	electron state	double bond equivalents (DBE) range	elements allowed
CLIPS	1	1–15	even	−0.5 to 30	C, H, N, O, F, P, Si, S, Cl, Br, I
sCLIPS	1	4–6	even	−0.5 to 30	C, H, N, O, F, P, Si, S, Cl, Br, I

observed that a lower mass tolerance could be used than for external calibration. Allowed elements in the formulas generated were those commonly found in organic contaminants, such as pharmaceuticals, pesticides, and plastic additives. The number range for each element was set automatically according to the Seven Golden Rules.⁸ In some cases, certain parameters had to be individually tuned. Such changes included a higher mass tolerance window if the mass calibration was off. Cl, Br, and Si atoms were also withdrawn to reduce computing time when the spectra visually did not suggest the presence of such atoms. Finally, in a few cases, chemically aberrant formulas, that is, containing an unrealistic number of H, F, or P atoms, were eliminated from the list given by MassWorks. Determination of the error of spectral accuracy is detailed in the Supporting Information (section SI-2).

RESULTS AND DISCUSSION

Impact of Matrix on Spectral Accuracy and Formula Ranking on QqTOFMS Data. Mass accuracy for the 80 and 300 $\mu\text{g L}^{-1}$ samples analyzed in the QqTOFMS can be seen in Table 2. As expected, values for the QqTOFMS were generally lower or equal than 2 ppm. It was hypothesized that environmental matrices such as river water would have no impact on formula ranking and spectral accuracy for the QqTOFMS since the risk of peak overlaps between analytes and matrix interferences would be much reduced in high resolution mass analyzers. However, results showed that the river water extract (matrix) did have an impact, albeit minor, on rankings (Table 3) and spectral accuracies (Table 4) measured with the QqTOFMS especially for the compounds with molecular mass >350 Da. For example, with samples spiked at a concentration of 80 $\mu\text{g L}^{-1}$, JOS (neutral nominal mass = 827 Da) was ranked 25 ± 24 in the MeOH solution and 34 ± 33 in the matrix. The impact of the matrix was however less pronounced on OFL (361 Da) and FLX (309 Da) which were always among the top 5 possible formulas in both MeOH and the matrix. Interestingly, OFL had a slightly worse ranking in the pure solvent than in the river extract, that is, 3 ± 2 in MeOH and ranked second or first in the matrix. This result was interpreted as a consequence of the presence of matrix effects causing signal enhancement that will be explained later.

Spectral accuracies were generally not affected by the matrix and differences between MeOH and river extracts differed only by approximately 3 percentage points. The slight decrease in spectral accuracy for SMX could be explained by the presence of matrix interferences near the peaks of the isotope pattern which were included in the spectral accuracy calculation. These interferences can be corrected with the interference rejection parameters which are discussed in the Supporting Information (section SI-3). For OFL spiked at 80 $\mu\text{g L}^{-1}$, the same trend in the ranking was observed in the spectral accuracy: higher spectral accuracy in the matrix extract ($96.3 \pm 0.1\%$) than in MeOH solution ($93.2 \pm 0.1\%$). Such improvement in ranking

Table 2. Mass Accuracy in ppm for the Test Compounds in the Three Different Mass Analyzers Employed^a

compounds	QqQMS				QqTOFMS ($R_{\text{FWHM}} = 2.5\text{K}$)				QqOrbitrapMS ($R_{\text{FWHM}} = 70\text{K}$)				QqOrbitrapMS ($R_{\text{FWHM}} = 140\text{K}$)			
	MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix	
	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	
ATZ (215)	2 ± 1	1.5 ± 0.3	5 ± 3	1.6 ± 0.8	2 ± 1	2 ± 2	3 ± 2	3 ± 1	2 ± 2	2 ± 2	3 ± 1	1 ± 1	2 ± 1	1.2 ± 0.9	3.3 ± 0.3	
SMX (253)	2 ± 2	1.0 ± 0.5	1.2 ± 0.7	1.1	0.7 ± 0.4	0.3 ± 0.2	1.0 ± 0.2	2.3 ± 0.2	3.3 ± 0.2	3.3 ± 0.2	2.3 ± 0.2	0.3 ± 0.2	0.5 ± 0.2	1.8 ± 0.2	0.4 ± 0.4	
MTP (267)	NA	0.2 ± 0.2	NA	1.1 ± 0.6	0.9 ± 0.2	0.6 ± 0.2	1.8 ± 0.9	1.0 ± 0.4	0.7 ± 0.4	1.0 ± 0.4	1.0 ± 0.4	0.3 ± 0.2	1.1 ± 0.2	1.0	0.9 ± 0.2	
TRI (290)	2 ± 2	0.3 ± 0.1	4 ± 3	0.7 ± 0.4	1 ± 1	0.8 ± 0.4	2.1 ± 0.7	0.6 ± 0.2	1.1 ± 0.5	0.6 ± 0.2	1.1 ± 0.5	0.3 ± 0.2	1.6 ± 0.3	0.6 ± 0.2	1.1 ± 0.2	
FLX (309)	NA	6.0 ± 0.9	NA	2 ± 2	2 ± 2	1.6 ± 0.2	2.6 ± 0.5	2 ± 2	3 ± 1	1.6 ± 0.2	2.6 ± 0.5	0.73	1.4	0.7 ± 0.9	1.5 ± 0.4	
OFL (361)	NA	1.5 ± 0.8	NA	0.70 ± 0.03	0.2 ± 0.2	1.2 ± 0.3	0.7 ± 0.7	1.4 ± 0.2	1.1 ± 0.2	1.2 ± 0.2	1.4 ± 0.2	1.4 ± 0.2	0.67	1.1 ± 0.3	0.9 ± 0.2	
MTX (454)	3 ± 2	1.0 ± 0.9	4 ± 3	2 ± 1	0.2 ± 0.2	0.8 ± 0.1	0.5 ± 0.1	1.4	0.5 ± 0.2	0.9 ± 0.2	0.9 ± 0.2	0.9 ± 0.2	0.2 ± 0.2	1.3 ± 0.3	0.5 ± 0.2	
JOS (827)	NA	1.2 ± 0.3	NA	1 ± 1	0.6 ± 0.3	2.3 ± 0.9	3.9 ± 0.8	1.8 ± 0.3	2.7 ± 0.5	1.8 ± 0.3	2.7 ± 0.5	1.2 ± 0.5	2.2 ± 0.2	0.9 ± 0.7	2.3 ± 0.6	
ROX (837)	0.7 ± 0.2	0.5 ± 0.2	0.8 ± 0.1	1.1 ± 0.7	1.0 ± 0.1	1.2 ± 0.6	4.3 ± 0.2	0.7 ± 0.5	2.1 ± 0.2	0.7 ± 0.5	2.1 ± 0.2	1.0 ± 0.4	2.2 ± 0.7	0.9 ± 0.2	2.2 ± 0.3	

^aNA: Not available. Uncertainty for all values represents standard deviation of three replicate injections. When uncertainty is absent, standard deviation was zero.

Table 3. Formula Ranking Results for the Test Compounds in the Three Different Mass Analyzers Employed^a

compounds	QqQMS						QqTOFMS ($R_{\text{FHWM}} = 25\text{K}$)						QqOrbitrapMS ($R_{\text{FHWM}} = 70\text{K}$)						QqOrbitrapMS ($R_{\text{FHWM}} = 140\text{K}$)						
	MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix		
	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	
ATZ (215)	3 ± 2	1	3.3 ± 0.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
SMX (253)	127 ± 103	1	107 ± 66	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
MTP (267)	NA	1	NA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
TRI (290)	9 ± 8	1	47 ± 69	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
FLX (309)	NA	1 ± 1	NA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
OFL (361)	NA	3 ± 2	NA	2 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1
MTX (454)	218 ± 110	5 ± 3	1177 ± 858	1	8 ± 3	2 ± 1	7 ± 1	5 ± 2	9 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	
JOS (827)	NA	25 ± 24	NA	3 ± 1	34 ± 33	11 ± 12	1	1 ± 1	1 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	
ROX (837)	57 ± 15	2 ± 2	48 ± 17	2 ± 1	4 ± 2	2 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	1	1 ± 1	

^aNA: Not available. Uncertainty for all values represents standard deviation of three replicate injections. When uncertainty is absent, standard deviation was zero.

Table 4. Spectral Accuracy Results for the Test Compounds in the Three Different Mass Analyzers Employed^a

compounds	QqQMS						QqTOFMS ($R_{\text{FHWM}} = 25\text{K}$)						QqOrbitrapMS ($R_{\text{FHWM}} = 70\text{K}$)						QqOrbitrapMS ($R_{\text{FHWM}} = 140\text{K}$)					
	MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix		MeOH		matrix	
	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$	300 $\mu\text{g L}^{-1}$	80 $\mu\text{g L}^{-1}$
ATZ (215)	90 ± 2	91 ± 1	98.5 ± 0.5	98.9 ± 0.1	98.96 ± 0.04	98.8 ± 0.7	97.7 ± 0.2	97.27 ± 0.04	97.5 ± 0.4	97.6 ± 0.2	97.08 ± 0.05	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2	97.3 ± 0.2	95.8 ± 0.2
SMX (253)	86 ± 10	94 ± 4	97.0 ± 0.5	98.5 ± 0.7	97.5 ± 0.5	91 ± 6	85.7 ± 0.5	95.8 ± 0.2	98.2 ± 0.4	98.7 ± 0.1	88 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3	98 ± 1	95.8 ± 0.3
MTP (267)	NA	NA	98.45 ± 0.02	96 ± 1	98.5 ± 0.2	97 ± 5	98.8 ± 0.5	98.2 ± 0.04	99.55 ± 0.08	98.2 ± 0.6	98.3 ± 0.2	96.7 ± 0.8	98.88 ± 0.03	96.0 ± 0.3	98.88 ± 0.03	96.0 ± 0.3	98.88 ± 0.03	96.0 ± 0.3	98.88 ± 0.03	96.0 ± 0.3	98.88 ± 0.03	96.0 ± 0.3	98.88 ± 0.03	96.0 ± 0.3
TRI (290)	97 ± 1	97 ± 1	97.98 ± 0.09	96 ± 1	98.6 ± 0.2	99.3 ± 0.2	97.70 ± 0.08	98.0 ± 0.2	98.5 ± 0.3	97.3 ± 0.8	98.1 ± 0.1	95.0 ± 0.8	98.2 ± 0.4	95 ± 1	98.2 ± 0.4	95 ± 1	98.2 ± 0.4	95 ± 1	98.2 ± 0.4	95 ± 1	98.2 ± 0.4	95 ± 1	98.2 ± 0.4	95 ± 1
FLX (309)	NA	NA	97.0 ± 0.4	97.4 ± 0.2	99.1 ± 0.1	97 ± 3	95.2 ± 0.4	98.6 ± 0.1	99.60 ± 0.06	99.3 ± 0.2	98 ± 1	98.90 ± 0.05	99.5 ± 0.1	99.13 ± 0.06	99.5 ± 0.1	99.13 ± 0.06	99.5 ± 0.1	99.13 ± 0.06	99.5 ± 0.1	99.13 ± 0.06	99.5 ± 0.1	99.13 ± 0.06	99.5 ± 0.1	99.13 ± 0.06
OFL (361)	NA	NA	93.17 ± 0.08	99.32 ± 0.07	98.7 ± 0.2	96 ± 6	96.6 ± 0.4	98.07 ± 0.02	98.4 ± 0.5	98.0 ± 0.3	93 ± 1	97.42 ± 0.07	99.03 ± 0.02	95.4 ± 0.4	99.03 ± 0.02	95.4 ± 0.4	99.03 ± 0.02	95.4 ± 0.4	99.03 ± 0.02	95.4 ± 0.4	99.03 ± 0.02	95.4 ± 0.4	99.03 ± 0.02	95.4 ± 0.4
MTX (454)	82 ± 9	86 ± 5	89 ± 6	92.2 ± 0.2	93.3 ± 0.2	97.9 ± 0.5	97.4 ± 0.6	97.82 ± 0.07	96.8 ± 0.1	98.52 ± 0.06	96.7 ± 0.2	96.501 ± 0.009	97.1 ± 0.2	97.2 ± 0.1	97.1 ± 0.2	97.2 ± 0.1	97.1 ± 0.2	97.2 ± 0.1	97.1 ± 0.2	97.2 ± 0.1	97.1 ± 0.2	97.2 ± 0.1	97.1 ± 0.2	97.2 ± 0.1
JOS (827)	NA	NA	92.19 ± 0.06	99.2 ± 0.1	96.8 ± 0.3	97 ± 2	99.5 ± 0.1	99.5 ± 0.1	99.0 ± 0.2	99.2 ± 0.1	99.29 ± 0.07	99.53 ± 0.02	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1	99.5 ± 0.1
ROX (837)	93 ± 3	95 ± 2	93.3 ± 0.1	99.2 ± 0.1	98.1 ± 0.1	98 ± 1	99.46 ± 0.06	99.46 ± 0.02	99.3 ± 0.1	99.22 ± 0.08	99.4 ± 0.2	99.3 ± 0.1	99.4 ± 0.1	99.65 ± 0.04	99.4 ± 0.1	99.65 ± 0.04	99.4 ± 0.1	99.65 ± 0.04	99.4 ± 0.1	99.65 ± 0.04	99.4 ± 0.1	99.65 ± 0.04	99.4 ± 0.1	99.65 ± 0.04

^aNA: Not available. Uncertainty for all values represents standard deviation of three replicate injections. When uncertainty is absent, standard deviation was zero.

and spectral accuracy for OFL could be explained by signal enhancement caused by coeluting sample components, which has been reported previously^{17,18} and resulted in an improved signal-to-noise ratio. The peak area for OFL was indeed about 3 to 20 times higher in the matrix compared to the MeOH solution (Table SI-2). Matrix effects were quantified by calculating the ratio of peak areas for a test compound in the matrix and MeOH. Thus, for the results presented in Table SI-2, area ratios >1 indicate signal enhancement and area ratios <1 signal suppression. Signal enhancement occurs because matrix components that coelute with target compounds can improve access to the droplet surface during the electrospray ionization process and thus increase ionization efficiency.¹⁷ The opposite effect, known as signal suppression, was also observed during the experiments. Signal suppression can be significant (up to 90%) in some cases in other environmental waters such as hospital and wastewater influents.¹⁹ While these matrix effects can be hardly predicted, they could be reduced by modification in the sample preparation and the chromatographic separation.

Impact of Analyte Concentration on Spectral Accuracy and Formula Ranking on QqTOFMS Data. Results in Table 3 show that concentration (80 vs 300 $\mu\text{g L}^{-1}$ spiked in the river water matrix) was an important factor in deciding the correct formula rank for the larger compounds (>350 Da). For smaller compounds (<350 Da), the lower number of possible outcomes for a given accurate mass requires less spectral accuracy for correct formula identification. This was especially true for ATZ and SMX which have very recognizable features due to their Cl and S atoms, respectively. Spectral accuracy of MTX (454 Da) (Table 4) was the most affected by concentration, from 90.7% \pm 0.9% in the 80 $\mu\text{g L}^{-1}$ matrix solution to 97.9% \pm 0.5% in the 300 $\mu\text{g L}^{-1}$ matrix solution. For the other compounds, differences were <4 percentage points. This can be explained by a higher signal for the M+2 and M+3 peaks which allows for a more thorough comparison of the calibrated experimental and theoretical peaks in MassWorks. When peak intensities were <800 counts, major discrepancies are observed between the calibrated experimental and theoretical peaks. These differences result in much lower overall spectral accuracy and less confidence in formula determination. Since the signals for the other compounds were relatively higher than those of MTX in the 80 $\mu\text{g L}^{-1}$ matrix solution, the increase in concentration had a less noticeable effect. Finally, for SMX and OFL, standard deviations >5 percentage points in spectral accuracies were observed. SMX standard deviation was indeed almost 4 times larger in the 300 $\mu\text{g L}^{-1}$ solution compared to the 80 $\mu\text{g L}^{-1}$ solution. Such high variation is due to an outlier value. When the outlier replicate was removed, the variation in spectral accuracy was within 2 percentage points.

Impact of Mass Analyzer Resolution on Spectral Accuracy and Identification. *Triple Quadrupole Mass Spectrometer (QqQMS).* Measurements with QqQMS were done using both internal and external mass calibrations. Better results were obtained using internal mass calibration and are discussed here. Results obtained using external calibration are discussed in the Supporting Information (section SI-1).

Mass accuracy results obtained after post-acquisition internal calibration and determination of molecular formulas based on the CLIPS algorithm can be seen in Table 2. For all compounds, average mass accuracy was ≤ 5 ppm. As expected, better mass accuracy was obtained in both QqTOFS and

QqOrbitrapMS, but it is remarkable that a QqQMS can attain such low errors in measuring accurate masses. However, concentration had a major impact on both ranking and spectral accuracy on the QqQMS data. In fact, it was possible to obtain meaningful data only for the samples spiked at 300 $\mu\text{g L}^{-1}$ (Tables 3 and 4). In general, the matrix slightly increased spectral accuracy but rank was not significantly affected. As explained previously, the higher spectral accuracy in the river matrix could be explained by signal enhancement during the ESI process, in which the coeluting compounds improve ionization efficiency. As it can be seen in the Supporting Information Table SI-2 for QqQMS data, matrix/MeOH area ratios were between 1.2 and 2.8, indicating that signal enhancement due to the matrix, were observed for all compounds except for ROX (0.9). However, for TRI and MTX, lower rankings despite the matrix induced signal enhancement were observed. These results illustrate the two opposite effects of the matrix on spectral accuracy and formula ranking: (1) ions of coeluting matrix compounds with *overlapping* m/z values with the compound of interest may affect intensities of the isotopic pattern and thus lower spectral accuracy and ranking and (2) ions of coeluting matrix compounds of *different* m/z from the compound of interest may increase spectral accuracy and formula ranking by increasing the signal-to-noise ratio of the all the peaks of the isotopic pattern.

As shown in Table 3, obtained rankings were not good enough to allow acceptable certainty in formula determination in most cases. Therefore, these results suggest that the use of a QqQMS for the measurement of accurate masses and spectral accuracy of small organic molecules at low concentrations in complex matrices seems to be very challenging. Nevertheless, it could be very helpful for other routine applications that use simpler matrices and compounds at higher concentrations, e.g. monitoring and confirmation of organic synthesis products or impurity and degradation identification in pharmaceutical products.

Quadrupole-Orbital Ion Trap Mass Spectrometer (QqOrbitrapMS). Mass accuracy and ranking results at 80 and 300 $\mu\text{g L}^{-1}$ in MeOH and the matrix solution obtained with QqOrbitrapMS at $R_{\text{FWHM}} = 70$ K and 140 K are shown in Table 2 and 3, respectively. As expected, mass accuracy for the QqOrbitrapMS were generally lower or equal than 2 ppm. Concerning ranking, both low mass compounds (<350 Da) and high mass compounds (between 350 and 837 Da) were not affected by resolution considering that rankings were consistently in the top ten for both resolutions. Generally, lower spectral accuracy was observed for the low mass compounds compared to the high mass compounds, but it did not appear to be affected by resolution at lower concentration. For MTP, TRI and OFL in the matrix samples spiked at the higher concentration (300 $\mu\text{g L}^{-1}$), the largest drops in spectral accuracy, caused by an increase in resolution were between 2.2 to 2.7 percentage points. Those differences were statistically significant at the 95% confidence level according to the t test. Such a reduction in spectral accuracy is only about half of that reported by a previous study using a linear ion trap-orbital ion trap mass spectrometer,¹² where the spectral accuracy of polar organic compounds (masses between 639 to 1664 Da) was higher ($\geq 97\%$) at $R_{\text{FWHM}} = 7.5$ K than that obtained at $R_{\text{FWHM}} = 100$ K (<90%). According to the authors of that study, high resolution hinders high spectral accuracy in the orbitrap mass spectrometer. The authors explained those results as the consequence of a phenomenon called "isotope beating". That phenomenon results from destructive interference of signals having close m/z values, for example, isotopic peaks of

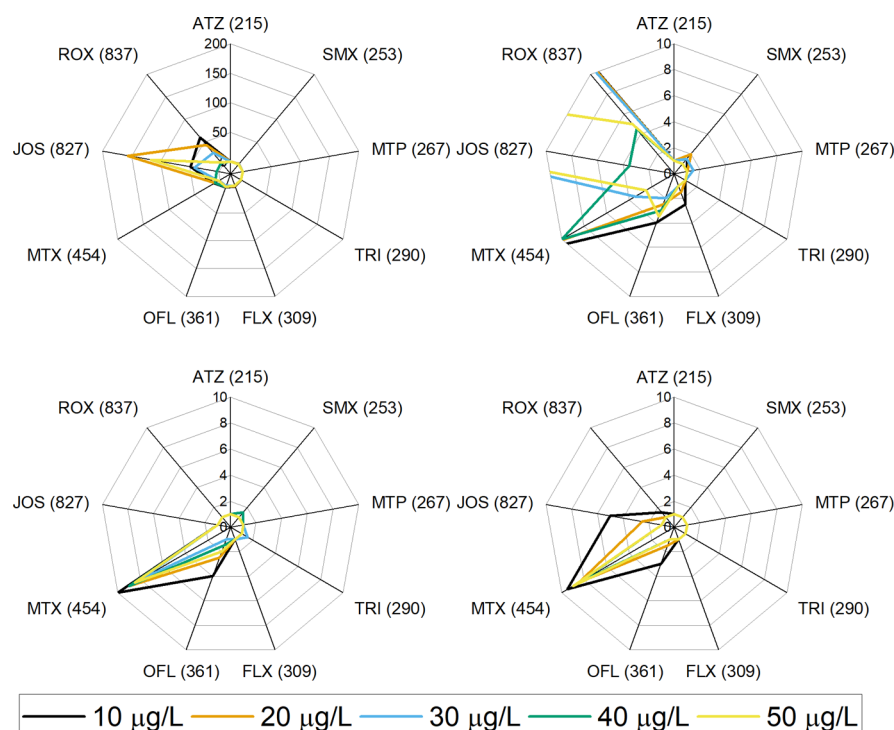


Figure 2. Radar plots representing mean formula ranking for the target micropollutants in spiked matrix samples. Top left: QqTOFMS ($R_{\text{FWHM}} = 25$ K). Top right: QqTOFMS results zoomed in. Bottom left: QqOrbitrapMS ($R_{\text{FWHM}} = 70$ K). Bottom right: QqOrbitrapMS ($R_{\text{FWHM}} = 140$ K).

multiply charged ions or closely located isotopes under a given isotope cluster, such as $M + 3$, and produce errors in the measurement of isotopic abundances.²⁰ Such effects have been observed with polymers in ion cyclotron resonance mass spectrometers (ICRMS)²⁰ and may also be present in the QqOrbitrapMS, which is also an ion trap mass analyzer using Fourier transform signal conversion. Interestingly, such negative correlation between resolution and spectral accuracy was not observed in the present study for compounds with the highest molecular masses (ROX, JOS). It is likely that such issues have been either reduced or corrected in the newer orbital ion trap models, especially at the high masses ($m/z > 800$). It is possible that enhanced Fourier transform for orbitrap mass spectrometry, introduced in 2014, brought some amelioration in peak shape with the addition of apodization and triple zero-filling. These notably helped reducing spectral leakage and side lobes around peaks.²⁰ Another possibility, proposed by Xu et al.,¹⁴ is the effect of the reduced-profile mode for the recording of mass spectra in orbitraps. In that mode, the noise is subtracted from the acquired mass spectra to reduce raw file size. This signal processing can lead to underestimation of the abundance of $M + n$ peaks, thus reducing spectral accuracy. It has yet to be confirmed how all these factors played a major role in reducing spectral accuracy with higher resolutions, and it is out of the scope of this article.

Although QqTOFMS are known to measure isotopic patterns very precisely⁸ and accurately¹³ compared to other mass analyzers, results showed that better spectral accuracies for the tested compounds were obtained with the QqOrbitrapMS at $R_{\text{FWHM}} = 70$ K and 140 K than with the QqTOFMS at $R_{\text{FWHM}} = 25$ K. As it can be seen in Table 4, in the data acquired with the QqTOFMS employing the matrix spiked at $300 \mu\text{g L}^{-1}$, only 3 out of 9 test compounds had average spectral accuracy $\geq 98\%$, while in the QqOrbitrapMS data ($R_{\text{FWHM}} = 70$ K), 7 out of 9 compounds had average spectral

accuracy $\geq 98\%$. Formula rankings were also better with the QqOrbitrapMS at both resolutions compared to the QqTOFMS.

It would be interesting to perform experiments at ultrahigh resolution with an ICRMS to see the effect of full separation of the isotope fine structure on spectral accuracy and formula ranking. Nevertheless, the sCLIPS algorithm does already factors in unresolved isotope fine structures in the calculation of mass accuracy since it takes into account instrument resolution to calibrate the mass spectrum signal.

An important observation from the QqOrbitrapMS results shown in Tables 2 and 3 is that low spectral accuracy (e.g., $< 98\%$) does not necessarily mean a low ranking as some compounds consistently get low spectral accuracies, in our case SMX data clearly illustrates that. It seldom or never had $> 98\%$ spectral accuracy but it was always ranked first nevertheless. Since there are less possible formulas for compounds with low molecular masses for a given accurate mass compared to compounds with high molecular masses, the former requires less spectral accuracy for correct formula identification than the latter.

Lower Concentration Limits for the Measurement of Spectral Accuracy in Environmental Samples.

Solutions of lower concentrations, ranging from 50 to $10 \mu\text{g/L}$, equivalent to environmental concentrations of 0.38 to $1.9 \mu\text{g L}^{-1}$ considering solid-phase extraction preconcentration factor, were analyzed with the intent to assess a minimum working concentration for spectral accuracy. Rankings and spectral accuracies can be seen in Figures 2 and 3, respectively. In the QqTOFMS, the larger compounds such as ROX and in particular JOS were heavily affected by the drop of signal intensity. JOS notably had very low signal intensity; thus $M+2$ peaks and onward were undistinguishable. Very high variability in the rankings was also observed in the cases of JOS (4–157) and, to a lesser degree, ROX (5–61). ROX did however improve significantly in ranking at 40 and $50 \mu\text{g L}^{-1}$. ROX signal intensity was also much superior to that of JOS (16 500 counts at apex at $50 \mu\text{g L}^{-1}$

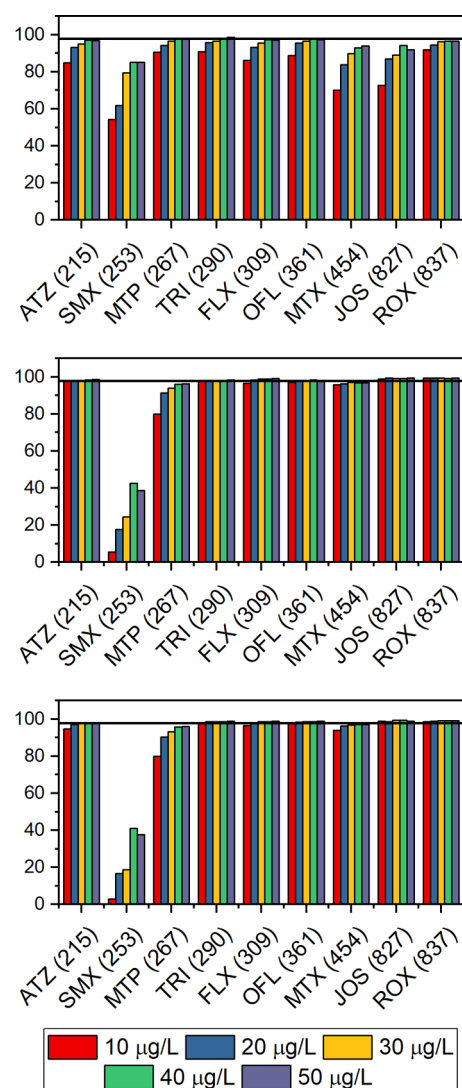


Figure 3. Bar plots representing mean spectral accuracy for the target micropollutants in spiked matrix samples. Top: QqTOFMS ($R_{\text{FWHM}} = 25$ K). Middle: QqOrbitrapMS ($R_{\text{FWHM}} = 70$ K). Bottom: QqOrbitrapMS ($R_{\text{FWHM}} = 140$ K). Straight line indicates the threshold of high spectral accuracy (98%). The same sample was injected twice.

compared to 2300 in the same conditions respectively); $M + 2$ and $M + 3$ were well-defined. Rankings were never lower than 4 for ATZ, SMX, MTP, TRI, FLX, and OFL in all instances. MTX ranking varied from 3 to 12.

Those lower concentrations mixtures were also analyzed with the QqOrbitrapMS at both 70 K and 140 K resolutions. Ranking and spectral accuracy results can be seen in Figures 2 and 3, respectively. Data showed excellent rankings (within top 5) for all compounds at all concentrations in both resolutions. Except for MTX, which ranked consistently over 5. OFL had higher ranking at the lower concentrations for both resolutions. JOS was also affected at 140K resolution for the 10 and 20 $\mu\text{g L}^{-1}$. Spectral accuracies were low for SMX, between 5.6% and 38.7% for 70 K. and between 3.0% and 37.7% for 140 K. This was caused by an impurity of high intensity at m/z 256, which although well separated from the $M + 2$ peak, lowered the match between calibrated and theoretical isotopic patterns of SMX. This impurity was not observed in the QqTOFMS data, thus spectral accuracy for SMX on that instrument was superior (54.2–85.1%) than in the QqOrbitrapMS.

As discussed in the previous sections, adequate measurement of spectral accuracy depends on many factors such as nature of the compound, matrix composition and even the chromatographic and sample preparation methods employed. Therefore, the lower concentration limits for the measurement of spectral accuracy can be improved by using more selective solid-phase extraction or by improving the chromatographic separation of analytes from coeluting matrix compounds having overlapping m/z values with the compound of interest. In summary, as long as the isotopic pattern is free of interferences and significantly higher than the noise, measurement of spectral accuracy will yield a dependable value that could be used to improve the level of confidence in the assignment of a molecular formula to an accurate mass.

Rules and Limitations. Comparison of spectral accuracy to other techniques used for formula determination showed that the algorithms used by MassWorks obtained better ranking of the correct formula of larger compounds (>350 Da), and it was more robust when using lower intensity signals (Supporting Information, section SI-4). However, this spectral accuracy determination has also a few drawbacks. Computing time necessary for determination of ranking and spectral accuracy of the high mass compounds such as JOS (827 Da) and ROX (837 Da) was the biggest downside found in this technique. Depending on the computer performance, the time to generate thousands of formula candidates could go up to 40 min if limitations were not made. Atoms like F and P being monoisotopic means they can be fitted in most formulas while a compound with high $M+1$ relative abundance like Si compensated for the lack of $M+1$ distribution from the inclusion of F and P. This results in a high number of generated formulas with low C and H and high Si and P, which are not realistic. In cases where a broad range of atoms are allowed in the formulas, pre- and post-research rules need to be defined to optimize formula generation using MassWorks. First, one must observe the mass spectra and visually assess the presence or absence of Cl and Br in the compounds. These two atoms are very distinct with unmistakable isotope patterns and they could be withdrawn from the allowed atom list if abundant $M + 2$, $M + 4$, or higher isotopes are not observed. Such procedure will divide by two the total number of generated formulas and thus save computing time. Additionally, the maximum number of C for a formula using the empirical parameters (Seven Golden Rules) needs to be used with utmost caution since it may exclude the correct formula from consideration. It was observed in one case that the empirical parameters on MassWorks underestimated the maximum number of C on a formula and the correct formula did not appear in the list. Such event was observed with a background contaminant, the plasticizer diisooctyl phthalate ($\text{C}_{24}\text{H}_{38}\text{O}_4$), where the maximum number of C was 23 based on the empirical rules implemented while the compound has 24 C. Setting a more reasonable minimum number of C also helps to reduce the number of generated formulas and saves computing time. A corresponding number of H can be added for the minimum and maximum values. A maximum H/C ratio of 2 was found to help reduce the number of formulas. P tends to be inserted in all formulas and can be monitored by a O/P ratio of minimum 3 as it mainly occurs in the form of organophosphates with high O per P; double bond equivalents (DBE) then also should be monitored as a P–O double bond implies a higher DBE. Na or K adducts were not observed for the selected compounds in this work, but adding them to the allowed atoms might help to uncover the accurate formula in

case positive alkaline adducts are formed. Adjusting pre-search parameters is not necessary with smaller molecules as there are only few possibilities for a compound with a narrow mass error window. Finally, lack of automation is also an issue as each compound mass spectrum needs to be evaluated individually.

CONCLUSION

This study showed that spectral accuracy is a powerful tool for the determination of chemical formulas from accurate mass data. Spectral accuracy allowed to reduce the number of likely molecular formulas for organic micropollutants of relative high molecular mass (e.g., between 400 and 900 Da) to less than 10, and in some cases, it assigned unambiguously one specific molecular formula to an experimental isotopic pattern. Experiments showed that the major parameter affecting spectral accuracy and correct formula ranking for a set of common organic micropollutants is signal intensity. Thus, conditions increasing signal intensity, such as signal enhancement by the matrix and higher compound concentration, favored higher spectral accuracy and ranking of the correct molecular formula. A significant improvement of both ranking and spectral accuracy was also obtained with higher mass resolution of the mass analyzer. Contrary to a previous study,¹³ a moderate (~7 percentage points) decrease in spectral accuracy with higher resolution in the orbitrap mass spectrometer was not observed.

Results also showed that high spectral accuracy (e.g., >98%) and identification of the correct molecular formula were not necessarily correlated for low molecular mass compounds (<350 Da). It was however more prevalent for high molecular mass compounds (>350 Da). Using MassWorks software it was possible to acquire accurate mass data with less than 5 ppm mass accuracy in a QqQMS. While the low resolution of the QqQMS impairs accurate mass and spectral accuracy determination in complex matrices such as river water, application of spectral accuracy to routine analyses is of interest for laboratories without access to high resolution MS technology. Experiments demonstrated that for some compounds, high spectral accuracies and rankings can be obtained at concentrations as low as 10 $\mu\text{g L}^{-1}$ and in general, if the isotopic pattern of the compound is free of major interferences and the signal is above the noise of the instrument, it is possible to measure spectral accuracy correctly.

Finally, this study confirmed that spectral accuracy could be used as a complementary technique to eliminate formula candidates corresponding to an observed accurate mass during identification workflows of organic micropollutants based on liquid chromatography-high resolution MS. Thus, spectral accuracy is a powerful tool to elevate level 5 data (accurate mass) to level 4 (unequivocal molecular formula) according to the identification confidence levels proposed by Schymanski et al.² In this way, identification of unknowns present in environmental samples can be a more efficient process.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b01761.

Molecular structures of the test compounds used in this study, external calibration and stability of the QqQMS system, error in spectral accuracy determination, interference rejection, determination of matrix effects in data acquired the three mass spectrometers, and software comparison (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel: 1-(819) 821-7922. Fax: 1-(819) 821-8019. E-mail: pa.segura@usherbrooke.ca.

ORCID

Pedro A. Segura: 0000-0002-4552-4903

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery grant awarded to P.A.S. We would like to thank Yongdong Wang and Don Kuehl of Cerno Biosciences for his help with MassWorks software. Funding sources did not have any involvement in the design, experiments, data interpretation, writing, revision, or submission of this study. We would like to thank Thermo Fisher Scientific for providing a generous access to a QqOrbitrapMS.

REFERENCES

- (1) Boxall, A.; Rudd, M.; Brooks, B.; Caldwell, D.; Choi, K.; Hickmann, S.; Innes, E.; Ostapyk, K.; Staveley, J.; Verslycke, T.; Ankley, G.; Beazley, K.; Belanger, S.; Berninger, J.; Carriquiriborde, P.; Coors, A.; Deleo, P.; Dyer, S.; Ericson, J.; Gagné, F.; et al. *Environ. Health Perspect.* **2012**, *120*, 1221–1229.
- (2) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098.
- (3) Gross, J. H. *Mass Spectrometry: A Textbook*, 2nd ed.; Springer: Berlin, Germany, 2011; p 716.
- (4) Eysseric, E.; Bellerose, X.; Lavoie, J.-M.; Segura, P. A. *Can. J. Chem.* **2016**, *94*, 781–787.
- (5) Yuan, M.; Breitkopf, S. B.; Yang, X.; Asara, J. M. *Nat. Protoc.* **2012**, *7*, 872.
- (6) Zhang, Y.; Ren, Y.; Jiao, J.; Li, D.; Zhang, Y. *Anal. Chem.* **2011**, *83*, 3297–3304.
- (7) Cortés-Francisco, N.; Flores, C.; Moyano, E.; Caixach, J. *Anal. Bioanal. Chem.* **2011**, *400*, 3595–3606.
- (8) Kind, T.; Fiehn, O. *BMC Bioinf.* **2007**, *8*, 105.
- (9) García-Reyes, J. F.; Hernando, M. D.; Molina-Díaz, A.; Fernández-Alba, A. R. *TrAC, Trends Anal. Chem.* **2007**, *26*, 828–841.
- (10) Little, J. L.; Cleven, C. D.; Brown, S. D. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 348–359.
- (11) Little, J. L.; Williams, A. J.; Pshenichnov, A.; Tkachenko, V. J. *Am. Soc. Mass Spectrom.* **2012**, *23*, 179–185.
- (12) Wang, Y.; Gu, M. *Anal. Chem.* **2010**, *82*, 7055–7062.
- (13) Erve, J. C.; Gu, M.; Wang, Y.; DeMaio, W.; Talaat, R. E. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2058–2069.
- (14) Xu, Y.; Heiliger, J.-F.; Madalinski, G.; Genin, E.; Ezan, E.; Tabet, J.-C.; Junot, C. *Anal. Chem.* **2010**, *82*, 5490–5501.
- (15) Hughes, S. R.; Kay, P.; Brown, L. E. *Environ. Sci. Technol.* **2013**, *47*, 661–677.
- (16) Segura, P. A.; Takada, H.; Correa, J. A.; El Saadi, K.; Koike, T.; Onwona-Agyeman, S.; Ofosu-Anim, J.; Sabi, E. B.; Wasonga, O. V.; Mghalu, J. M.; dos Santos, A. M., Jr.; Newman, B.; Weerts, S.; Yargeau, V. *Environ. Int.* **2015**, *80*, 89–97.
- (17) Taylor, P. J. *Clin. Biochem.* **2005**, *38*, 328–334.
- (18) Segura, P. A.; Gagnon, C.; Sauvé, S. *Anal. Chim. Acta* **2007**, *604*, 147–157.
- (19) Gros, M.; Rodríguez-Mozaz, S.; Barceló, D. *J. Chromatogr. A* **2013**, *1292*, 173–188.
- (20) Lange, O.; Damoc, E.; Wiegand, A.; Makarov, A. *Int. J. Mass Spectrom.* **2014**, *369*, 16–22.